

Read the following article and then answer the questions at the end.

The Age of Intelligent Machines: The Social Impact of Artificial Intelligence by Margaret A. Boden

Is artificial intelligence in human society a utopian dream or a Faustian nightmare? Will our descendants honor us for making machines do things that human minds do or berate us for irresponsibility and hubris? In this chapter from *The Age of Intelligent Machines* (published in 1990), Margaret A. Boden explores the potential impacts of artificial intelligence on society.

Is artificial intelligence in human society a utopian dream or a Faustian nightmare? Will our descendants honor us for making machines do things that human minds do or berate us for irresponsibility and hubris? Either of these judgments might be made of us, for like most human projects this infant technology is ambivalent. Just which aspects of its potential are realized will depend largely on social and political factors. Although these are not wholly subject to deliberate control, they can be influenced by human choice and public opinion. If future generations are to have reason to thank us rather than to curse us, it's important that the public (and politicians) of today should know as much as possible about the potential effects-for good or ill-of artificial intelligence (AI).

What are some of the potential advantages of AI? Clearly, AI can make knowledge more widely available. We shall certainly see a wide variety of expert systems: for aiding medical diagnosis and prescription, for helping scientists, lawyers, welfare advisers, and other professionals, and for providing people with information and suggestions for solving problems in the privacy of their homes. Educational expert systems include interactive programs that can help students (schoolchildren or adults, such as medical students) to familiarize themselves with some established domain. This would give us much more than a set of useful tools and educational cribs. In virtue of its applications in the communication and exploration of knowledge, AI could revolutionize our capacity for creativity and problem solving, much as the invention of printing did.

One advantage of having computers in the schoolroom and elsewhere is that they are not human. Precisely because they are not, they will not be bored by their human user's questions, nor scorn their user's mistakes, as another person might. The user may be ignorant, stupid, or naive, but the computer will not think so. Moreover, what looks like ignorance, stupidity, or naiveté is often a sort of exploratory playing around with ideas that is the essence of learning and of creativity. Many children have their self-confidence undermined by their teacher's explicit or implicit rejection of their attempts at self-directed thinking. Similarly, many people-for instance, those who are - female, working class, Jewish, disabled, or black-encounter unspoken (and often unconscious) prejudice in their dealings with official or professional bodies. An AI welfare adviser, for example, would not be prejudiced against such clients unless its data and inferential rules were biased in the relevant ways. A program could, of course, be written so as to embody its programmer's prejudices, but the program can be printed out and examined, whereas social attitudes cannot.

Artificial intelligence might even lead to a society in which people have greater freedom and greater incentive to concentrate on what is most fully human. Too few of us today (especially men) have time to commit ourselves to developing our interpersonal relations with family and friends. Increased leisure time throughout society (on the assumption that appropriate political and economic structures had been developed to allow for this) would make room for such conviviality. Partly as a result of this and perhaps partly as a reaction against the unemotional nature of most AI programs, the emotional dimension of personality might come to be more highly valued (again, especially by men) than it is in the West today. In my view, this would be all to the good. Similarly, the new technology might make it possible for many more people (yet again, especially men) to engage in activities, whether paid or unpaid, in the service sector: education, health, recreation, and welfare. The need for such activities is pressing, but the current distribution of income makes these intrinsically satisfying jobs financially unattractive. One of the most important benefits of all is that AI can rehumanize-yes, rehumanize-our image of ourselves. How can this be? Most people assume that AI either has nothing to teach us about the nature of being human or that it depicts us as "nothing but machines": poor deluded folk, we believe ourselves to be purposive, responsible creatures whereas in reality we are nothing of the kind.

The crucial point is that AI is concerned with representations, and how they can be constructed, stored, accessed, compared, and transformed. A computer program is itself a set of representations, a symbol system that models the world more or less adequately. This is why it is possible for an AI program to reflect the sexist or racist prejudices of its programmer. But representation is central to psychology as well, for the mind too is a system that represents the world and possible worlds in various ways. Our hopes, fears, beliefs, memories, perceptions, intentions, and desires all involve our ideas about (our mental models of) the world and other worlds. This is what humanist philosophers and psychologists have always said, of course, but until recently they had no support from science. Because sciences like physics and chemistry have no place for the concept of representation, their philosophical influence over the past four centuries has been insidiously dehumanizing. The mechanization of our world picture-including our image of man was inevitable, for what a science cannot describe it cannot recognize. Not only can artificial intelligence recognize the mind (as distinct from the body); it can also help to explain it. It "gives us back to ourselves," by helping us to understand how it is possible for a representational system to be embodied in a physical mechanism (brain or computer).

So much for the rose-colored spectacles. What of the darker implications? Many people fear that in developing AI, we may be sowing the seeds of our own destruction, our own physical, political, economical, and moral destruction. Physical destruction could conceivably result from the current plans to use AI within the U.S. Strategic Defense Initiative (Star Wars). One highly respected computer scientist, David Parnas, publicly resigned from the U.S. government's top advisory committee on SDI computing on the grounds that computer technology (and AI in particular) cannot in principle achieve the reliability required for a use where even one failure could be disastrous. Having worked on military applications throughout his professional life, Parnas had no political ax to grind. His resignation, like his testimony before the U.S. Senate in December 1985, was based on purely technical judgment.

Political destruction could result from the exploitation of AI (and highly centralized telecommunications) by a totalitarian state. If AI research had developed programs with a capacity for understanding text, understanding speech, interpreting images, and updating memory, the amount of information about individuals that was potentially available to government would be enormous. Good news for Big Brother, perhaps, but not for you and me.

Economic destruction might happen too if changes in the patterns and/or rates of employment are not accompanied by radical structural changes in industrial society and in the way people think about work. Economists differ about whether the convivial society described above is even possible: some argue that no stable economic system could exist in which only a small fraction of the people do productive (nonservice) work. Certainly, if anything like this is to be achieved, and achieved without horrendous social costs, new ways of defining and distributing society's goods will have to be found. At the same time, our notion of work will have to change: the Protestant ethic is not appropriate for a high-technology postindustrial society.

Last, what of moral destruction: could we become less human-indeed, less than human-as a result of advances in AI? This might happen if people were to come to believe that purpose, choice, hope, and responsibility are all sentimental illusions. Those who believe that they have no choice, no autonomy, are unlikely to try to exercise it. But this need not happen, for our goals and beliefs-in a word, our subjectivity-are not threatened by AI. As we have seen, the philosophical implications of AI are the reverse of what they are commonly assumed to be: properly understood, AI is not dehumanizing.

A practical corollary of this apparently abstract point is that we must not abandon our responsibility for evaluating-and, if necessary, rejecting-the "advice" or "conclusions" of computer programs. Precisely because a program is a symbolic representation of the world, rather than a part of the world objectively considered, it is in principle open to question. A program functions in virtue of its data, its inferential rules, and its values (decision criteria), each and every one of which may be inadequate in various ways. (Think of the example of the racist expert system.) We take it for granted that human beings, including experts (perhaps especially experts), can be mistaken or ill advised about any of these three aspects of thinking. We must equally take it for granted that computer programs - which in any event are far less subtle and commonsensical than their programmers and even their users - can be questioned too. If we ever forget that "It's true because the computer says so" is never adequate justification, the social impact of AI will be horrendous indeed.

From The Age of Intelligent Machines, 1990

Questions:

If in the future machines have the ability to reason, be self-aware and have feelings, then what makes a human being a human being, and a robot a robot?

If you could have a robot that would do any task you like, a companion to do all the work that you prefer not to, would you? And if so, how do you think this might affect you as a person?

Are there any kind of robots that shouldn't be created? Or that you wouldn't want to see created? Why?

Automation and the development of new technologies like robots is viewed by most people as inevitable. But many workers who lose their jobs consider this business practice unfair. Do you think the development of new technologies, and their implementation, is inevitable? What, if anything, should we as a society do for those people who lose their jobs?